**Exercise 7:** Gene finding

The **ORF Finder** (Open Reading Frame Finder) is a graphical analysis tool which finds all open reading frames of a selectable minimum size in a user's sequence or in a sequence already in the database.

* Open [ORF finder](http://www.ncbi.nlm.nih.gov/gorf/gorf.html)
* enter the accession number "NM\_004006" (Dystrophin transcript variant Dp427m) into the appropriate box.

You will see this full cDNA sequence translated in all 6 reading frames with the results shown in diagrammatic format. Remember that DNA is double stranded so there are 3 reading frames for the top strand and 3 more reading frames for the bottom strand. All the ORFs are indicated by the colored boxes with the rest of the cDNA left uncolored. The ORF Finder is a graphical-analysis tool which allows you to find all open reading frames in a given sequence

|  |  |
| --- | --- |
|  | Open Reading Frame from the list |
|  | Selected Open Reading Frame |
|  | Accepted Open Reading Frame |

What you are generally looking for is the largest ORF. The software has already done this for you by ranking the ORFs from largest to smallest. It also tells you which frame it came from, for example +3 means it is the third frame (3) on the top strand (+).

* Click on the colored box next to the +2 from the list on the largest ORF. You will see the color of the ORF box change and the deduced amino acid sequence.
* "Accept" the proposed ORF and View Genbank entry and find the Gene name